WRITTEN STATEMENT OF
RON S. JARMIN
ACTING DIRECTOR, U.S. CENSUS BUREAU
U.S. DEPARTMENT OF COMMERCE

ON
"THE 2020 CENSUS AND CURRENT ACTIVITIES OF THE U.S. CENSUS BUREAU"

BEFORE THE
SENATE COMMITTEE ON HOMELAND SECURITY AND GOVERNMENTAL AFFAIRS
MARCH 23, 2021

Good afternoon Chairman Peters, Ranking Member Portman and members of the committee. It is an honor to be here today, and to highlight the U.S. Census Bureau's work. I appreciate the opportunity to talk to you today about the successes and challenges of the 2020 Census, how the pandemic spurred innovation in our non-decennial data and what lies ahead of us.

It was a challenging year for all of us, including at the Census Bureau. However, the 2020 Census used new tools to provide a snapshot of our nation—who we are, where we live, and so much more. We successfully offered an online response option which did not experience any downtime or hacks. A new phone option also provided non-English and non-Spanish language speakers an alternative to providing a response by mail. We ran the nation's largest advertising campaign, ultimately advertising in 46 languages, making new commercials to reflect the new realities, and finding new ways to support partners, including Members of Congress who helped us encourage their constituents to respond. We created new surveys that helped the country get valuable information to respond to the pandemic and learned to do things from home we never thought possible. And through it all we did everything we could to keep the public and our employees safe during the pandemic.

We took a big step forward in the use of technology for the 2020 Census and it paid off more than we anticipated. Our original objective was to modernize the decennial census – to make it more efficient and accessible to the public, and to help the Bureau more strategically deploy its resources. But those innovations also allowed us to adapt to the challenges of the pandemic, hurricanes, and wildfires to successfully complete data collection for the 2020 Census. Through it all, your support was invaluable, and I thank you for your oversight and your recognition of our work's value. We are working with outside experts to ensure that there is an independent assessment of 2020 Census data quality and implementing new ways of protecting people's individual responses and information. We are on track to produce the apportionment counts by April 30, 2021, and the redistricting data by September 30, 2021.

At the beginning of each decade, we try to design a census that will thrive in ten years not knowing entirely what things will look like ten years in the future. We spend the next several years building systems and processes that are strong enough and adaptable enough to withstand whatever that future world brings. We design. We build. We test. We refine. Repeatedly. We

have only one shot to count every person living in the country and we want the count to be as complete and accurate as possible. We understand the great responsibility that comes with producing data that shapes every person's future and every community's future for the next decade.

The Bureau's ability to innovate and respond to the challenges of the COVID-19 pandemic was not limited to 2020 Census operations. In the last year, we created new surveys and models - designing and implementing them faster than we ever have before – to examine the impact of the pandemic on businesses and households. We also produced tools and information for families, communities and decision makers who were responding to the pandemic.

Throughout all of this, we kept our regular survey work going, largely from home, while we mourned the loss of family members, friends and colleagues from COVID-19, juggled childcare and demanding work, and watched the world change in numerous ways. I have never been prouder to work alongside the talented public servants at the Census Bureau and I am privileged to showcase our work today.

2020 Census Pandemic Adaptations

Last March, the online response website opened and households started receiving invitations to respond to the 2020 Census – just as many parts of the country implemented stay-at-home orders. At the Census Bureau, the last time in the office for many of us was only shortly before Census Day. Initially, we suspended all in-person work on the census including our Update Leave operation, which delivers invitations to complete the census to households in largely rural areas that don't receive mail at their home or have recently been impacted by natural disasters. We also paused hiring census takers. During that time, the 2020 Census innovations were instrumental. The newly introduced online response option complemented mail response as did our new phone data collection option. We also began to train new staff virtually. We quickly adjusted and adapted our operations on many other fronts.

Throughout the summer and early fall, we made daily decisions on which operations to resume or start, and where and when to do so safely. The decisions were data-driven, using real-time information from sources like the Centers for Disease Control and Prevention, state, local and tribal governments and the National Weather Service. We also had to contend with numerous severe weather events, such as wildfires and hurricanes.

Guiding all of our decisions was our goal to conduct a complete and accurate count, while keeping the public and our employees safe. We tried to be as creative as possible and not limit ourselves just to things we had done before. To do this, the Census Bureau implemented new ways to respond and adjusted our operations to keep motivating the public and collecting responses. For example, we:

- Deployed staff to areas with lower response rates to answer people's questions and help them respond;
- Called households that had not responded instead of, or in addition to, sending a census taker;

- Emailed households in low responding areas to encourage them to respond online;
- Extended the time that the mail, online and phone options were available by two and a half months;
- Worked with local tribal leaders to hire individuals already living on reservation lands as census takers since their borders were closed due to the pandemic;
- Adapted our telephone operations to allow social distancing in the call centers. We also added a feature so callers could request that we call them back when wait times were long due to pandemic-related changes;
- Sent teams of skilled census takers from high response rate areas to locations lagging behind after hurricane damage;
- Changed our procedures to have census takers leave invitations on mailboxes, wear masks and remain socially distant;
- Provided and encouraged electronic response options for group quarters facilities, such as nursing homes and detention centers;
- Made and released new commercials that spoke to our pandemic environment before most other advertisers. And we got creative. Since people were not seeing ads on the street or in public places, we bought ads on pizza boxes because pizza deliveries soared. We bought advertising on new TV shows because people were watching TV during the day. We also placed advertising in the few places people were going like grocery stores and pharmacies;
- Increased the number of languages from 13 to 46 so we could reach more people;
- Held virtual concerts showcasing the benefits of responding to the census;
- Created new educational materials through our Statistics in Schools program that caregivers could use with children who were learning from home; and,
- Expanded outreach through our 400,000-plus national and local partners to encourage the public to respond or to cooperate with census takers.

The value provided by stakeholders and partners around the country who dedicated themselves to educating the public about the census and motivating them to respond was enormous. They also had to find new ways to carry their trusted messages into communities and we had to support them in fresh ways.

Although some in-person events were held before the onset of the pandemic—both to engage partners and to reach their audiences—most events scheduled were canceled. Virtual and digital channels were part of our original outreach plans, but these channels quickly became the primary means of engagement and activation. Channels included:

- Virtual events and webinars;
- Information dissemination to umbrella and trade organizations that represented numerous national organizations;
- Our website, which hosted numerous materials for stakeholders and partners to use;
- Social media outreach conducted through Census Bureau accounts; and,
- News releases, media tours, and earned media coverage.

We relied on our website and email to share materials and resources with partners during the pandemic. We shipped promotional items to stakeholders and partners that were included in meal kits and care packages distributed to help communities during the pandemic. Partners shared 2020 Census materials and information with community members through their websites, social media accounts, email listservs, and text message platforms. These channels were especially useful paths to the public since the use of social and digital media increased significantly during the pandemic.

Partners and stakeholders have always been critical in helping us reach historically undercounted groups, like people of color, renters and young children, but that was especially true in the COVID-19 environment, which has been marked by increased public anxiety and a busy media landscape. Because of the success of the 2020 Census partnership efforts in these communities, some advocates are now looking to leverage the relationships built at local and national levels between partners and stakeholders to promote vaccination.

Response Numbers

A pandemic – and all of the other events of the year from wildfires to hurricanes to other circumstances – could have easily depressed response rates. I am pleased to report that two in every three households responded on their own (67 percent), surpassing the 2010 Census response rate (66.5 percent). Of those households, 80 percent did so online. Fortunately, we did not experience one single minute of downtime or any cyber intrusions to our response website. We collected paper responses (18 percent) and phone responses (2 percent) as well, the latter of which was new for the 2020 Census. And finally, we visited remaining households in person to collect responses.

As a result of our extraordinary efforts, we were able to account for over 99.9 percent of the more than 152 million addresses in the nation. While we are proud of the completion rate, we know that's only part of the story and we have more to do to assess the quality of the data. We are eager to see how well we counted the people within those addresses and we are working to measure that now as we conduct the Post-Enumeration Survey. We are also working with independent, external experts, which I will discuss later.

Data Processing and Anomalies

Once we finished collecting data on October 15, 2021, we began processing it. Processing a census is complex work that takes time, computing power, and subject matter expertise from our staff – many of whom spend their entire careers doing this type of work. There are multiple required processing phases and, with each one, we are rigorously reviewing the resulting data files to ensure the 2020 Census counts are as accurate and complete as possible.

During data processing, we have discovered some "anomalies" along the way, as any data collection organization does in any census and survey. To be clear, anomalies are not errors in the census, but they can turn into errors if we do not review and resolve them. It is a feature of

our quality check process to find them, and it gives us the opportunity to fix any issues we confirm. Finding and resolving these anomalies shows that our quality checks are working.

Our subject matter experts meticulously comb through the response data, comparing population totals against other data sources, such as the 2010 Census, the 2020 population estimates, and Census Bureau's ongoing American Community Survey. We review the data, we look for outliers and, where we find them, we dig deeper to find out what is going on. Our career professionals also ensure that data processing ran as designed. If we determine a fix is needed to correct an anomaly, we fix it. Examining outliers is a normal part of data processing and the quality checks we do for any census or survey.

To date for the 2020 Census we have encountered 33 anomalies, which fall into three main categories:

I. 27 standard (or coding-related) anomalies;
II. 5 anomalies from unanticipated respondent actions (or actions we did not expect the public to take); and,
III. 1 anomaly from unanticipated census taker actions (or actions we did not expect census takers to take).

Standard anomalies occur in processing any census or survey. They relate to coding – how the response data appear and are processed in our data files and in the resulting tallies. Fortunately. we have resolved every anomaly that our systems and processes have identified, and we will continue to look for and address any that arise as we continue processing the data.

It is our intention to produce the apportionment counts by April 30, 2021, and the redistricting data by September 30, 2021. In recognition of the difficulties this timeline creates for states with redistricting and election deadlines prior to September 30, we have reviewed our timeline to identify any opportunities to shorten the processing schedule. We can provide a legacy format summary redistricting data file to all states by mid-to-late August 2021. Because we recognize that most states lack the capacity or resources to tabulate the data from these summary files on their own, we reaffirm our commitment to providing all states tabulated data in our user-friendly system by September 30.

Data Quality

For the first time, we plan to release data quality metrics for the nation as a whole, along with specific data for each of the states, District of Columbia, and Puerto Rico, along with the first 2020 Census results this April. Data quality metrics will include information on how many people responded on their own and how many households were counted with census taker visits, as well as metrics on addresses that were marked as occupied, vacant, or nonexistent. These operational quality metrics give the public an unparalleled degree of transparency into our work, providing metrics that previously were published a year or more after the census. These metrics are in addition to early indicators the Census Bureau has already provided for the 2020 Census, including self-response rates down to the tract level, initial completion rates at the state level,

nonresponse followup workload completion rates for area census offices, and national administrative records and proxy respondent enumeration rates.

Knowing that the COVID-19 pandemic might pose data quality concerns for the 2020 Census, I established the 2020 Data Quality Executive Governance Group last April to ensure resources were adequately and appropriately focused on addressing and documenting data quality issues. We also are enlisting outside experts to examine our quality metrics and review our 2020 Census processes, procedures, and decisions. To date, we have asked two highly respected, independent groups to do this work as we are processing the data and are working to engage a third group that will help us look ahead to the 2030 Census.

The first of these groups, JASON, an independent group of scientists and engineers that performs studies for the government, recently published the results of their evaluation. They called for improved communications strategies around the 2020 Census and recommended that we allow adequate time for complete, accurate and transparent data processing. JASON also expressed support for our proposed data quality metrics and provided recommendations for additional analysis, such as assessments of data quality across various geographies and for demographic groups. Finally, JASON stressed the importance of leveraging our experience with the 2020 Census to inform 2030 Census planning. All of this work is already underway.

Second, experts from the American Statistical Association (ASA) Quality Indicators Task Force are already working closely with us to build on the JASON report. Unlike JASON, the ASA is diving into the internal operational and response data from the 2020 Census to independently assess its accuracy and coverage. This work will unfold over the next year and will move in real-time along with Census Bureau assessments of data quality.

Similar to the ASA, the National Academy of Sciences (NAS) Committee on National Statistics (CNSTAT) has worked with the Census Bureau each decade, and it too includes many experts familiar with our work. Over the decades, CNSTAT has established panels to assess each decennial census and to suggest parameters for the research and planning of the subsequent censuses. We are still planning how they will engage with the 2020 Census.

These three groups will tackle different aspects of assessing the Census Bureau's work. Their reports will advise the Census Bureau on improving future censuses and will help the public understand the 2020 Census's quality. We look forward to their findings.

Next Steps: Upcoming 2020 Census Results

If the 2020 Census were a typical decennial census, we would be on the verge of delivering the first round of redistricting data. But we made schedule changes to keep people safe during data collection and to make sure we can process the data as thoroughly and carefully as we always have. As a result, those changes mean that we are on track to deliver the apportionment results by April 30. For the first time ever, we also plan to release operational quality metrics at that time. We know that all of the pandemic-induced changes raised questions about data quality and

the timeline. We want to maximize transparency and make sure that everyone feels confident in the numbers we release.

Our plan is to deliver the redistricting data to the states and the public by September 30, 2021. This data delivery will be a single national delivery, rather than the staggered delivery we originally planned. This change ensures that we can provide data that meet the quality standards states expect in the least amount of total time for all states. Additionally, just last week, on March 15 we announced that we would release an interim redistricting data product in mid-to-late August. This product will be less user-friendly for states, but it will contain the same data – and will have gone through the same exacting quality reviews – as the September release.

Differential Privacy

Like most federal agencies, the Department of Commerce and the Census Bureau have faced an increasingly difficult task in safeguarding privacy over the last few decades. Since the 1920s, we have used a variety of methods for protecting personal and business information. Our simplest early methods relied on suppression – manually hiding data points – and compression – combining data into larger categories.

Simple models for privacy protection became inadequate for protecting the privacy of data as computing power increased. In later decades, our methods had to adapt to this new reality, relying on suppressing whole tables in the 1970s and 80s, which threatened the usability of our data for many of the purposes we know are important to our data users. From 1990 through the 2010 Census, we began relying on more complex techniques for protecting personal information as computing power made the possibility of breaking through our safeguards more viable.

In the years leading up to the 2020 Census, we knew we needed an even more secure system for protecting privacy due to the sophisticated tools available to bad actors that enable them to identify individual respondents in our census and survey data. Having successfully used today's most sophisticated method for protecting privacy – called differential privacy – in another data set since 2008, we tested its applicability to decennial census data. After years of testing, we are confident that differential privacy will protect people's information while still ensuring that the statistics that we publish are useful. At a basic level, differential privacy is a privacy protection model that uses an algorithm – a mathematical formula – to strategically inject a level of uncertainty into our data. In contrast to privacy protection models we have used in the past that are comparatively blunt instruments for ensuring that we don't disclose an individual's information, differential privacy allows us to adjust the overall balance between privacy and accuracy and to 'tune' our algorithm to ensure we are meeting specific accuracy targets.

When releasing statistics based on data collected in the decennial census, the Census Bureau must guard against attempts to reidentify the people who responded. By using a highly evolved process of elimination, bad actors are able to reconstruct the source information – our survey responses – from the statistics we release.

Redistricting data released by the Census Bureau in the P.L. 94-171 file contain more than three billion data points. Additional demographic data released subsequently include an additional seven to ten billion pieces of information about the public. Armed with more than 10 billion extremely accurate statistical summaries, using today's computing power it is possible to accurately recreate the census responses of every one of the 330 million people and 140 million households in the country. Although we do not publish identifying information in those 10 billion statistical summaries, bad actors can easily link those reconstructed records to external databases containing the names and addresses of respondents. For the 2010 Census, our own researchers correctly identified more than 52 million respondents using these methods with only some of the available data. We would expect a bad actor utilizing all available data to identify even more information. Attacks like this get more sophisticated every day.

Privacy protection models are designed to help us guard against putting the personal information of 52 million or more people at risk. Yet the risk to the privacy of our respondents isn't our only concern when publishing data from the decennial census. We also need to ensure the data we're releasing are usable – that they are sufficiently accurate to meet the needs of our data users. Differential privacy allows us to strike a balance between privacy and accuracy in a surgical way.

The balance between privacy and accuracy has a numerical value in the differential privacy algorithm, called the privacy-loss budget. If we set that number very low, meaning we favor privacy, then when bad actors try to recreate the census responses to determine the age, sex, race, ethnicity, household size, location and so forth of each person who responded to the Census, they will almost certainly fail. Unfortunately, this would mean the data about our towns and neighborhoods would be so inaccurate that they are nearly meaningless. When we protect information with even a relatively small privacy-loss budget, statistics about our towns and neighborhoods become much more accurate because they are based on aggregating people and households, just as our simpler privacy protection models did. If the privacy-loss budget is set very high, meaning we favor accuracy, statistics about our towns and neighborhoods become nearly 'perfect.' Unfortunately, this would mean we would be putting the personal information of those 52 million or more at risk.

Finding the right privacy loss budget is only one step in designing the differential privacy system. In addition to setting this overall balance between privacy and accuracy, differential privacy lets us allocate that privacy-loss budget surgically to ensure we are meeting the needs of our data users. We can allocate our budget among geographies (like census blocks or municipal districts) and among characteristics (like race, ethnicity, and sex) to ensure that the accuracy of those values meets the needs of our users.

Because of how important our data is to decisions made every day, we know there is great interest in whether this new method of protecting individuals' information will change the utility of our statistics. We previously took 2010 data and applied differential privacy to it so users could provide feedback on changes to the algorithm as it was refined and so that they could better understand what impacts differential privacy would have when the model heavily favored privacy. Soon, we will release another set of example data that will use a better balance of

privacy and accuracy. We welcome conversations with stakeholders about where we should ultimately set that balance point so that we are protecting people's information while still producing useful statistics.

Other Pandemic Adaptations

While our decennial census team was busy adapting its operations and advertising campaign to the pandemic, other Census Bureau teams were quickly figuring out what data could help communities and governments around the country respond to the pandemic's impacts. The new data, estimates and tools we created allowed decision makers to understand the pandemic's effects on businesses and households by demonstrating which communities were most resilient to the health and economic challenges that arose and where the hardest hit communities and populations lived so response could be targeted equitably.

Before the pandemic hit, we released a set of statistics that examined how many new businesses were forming, how long it took them to form, and how quickly they were up and running to the point where they were paying employees. These numbers are great indicators of entrepreneurial activity, which can be stimulated or suffer in circumstances like a pandemic. Once the pandemic hit, we began releasing these numbers, called Business Formation Statistics, every week instead of every month, allowing for a near real-time snapshot of the impact on businesses. An advantage of this type of effort is that it uses existing information – applications for Employer Identification Numbers – so we are not asking people to take time in an extraordinary circumstance to provide information that allows us to measure the pandemic's impacts. Users can understand impacts by geography and industry with these numbers. Interesting information has come from this work. There was a sharp drop in business applications at the very start of the pandemic which was quickly followed by a surge in business applications that resulted in all-time highs. The surge in business applications since May 2020 is uneven across industries, but led by retail, in particular Nonstore Retailers, which includes much of online selling.

We also developed and implemented new surveys faster than ever before to give us a deeper understanding of the pandemic's impact. As evidenced by our decennial census planning and testing schedule, we are used to spending a good bit of time planning out our surveys and censuses, testing those plans, refining, and then executing. But in a matter of weeks, we planned and launched two experimental surveys - the Household and Small Business Pulse Surveys - to measure the pandemic's impacts. These surveys give businesses and households about a week or two, respectively, to respond. We publish business data every week and household data every other week. The surveys are designed so those responding can spend just a few minutes on them. Over the course of nearly a year of doing these surveys, we have worked with other federal agencies to make sure the questions we ask remain relevant to the circumstances households and businesses are experiencing.

The Small Business Pulse Survey, which focuses on businesses with no more than 499 employees, asks questions about businesses closing, supply chain disruptions, whether they are relying on federal assistance like the Paycheck Protection Program loans, whether they will require employees to be vaccinated, if a business has had to reduce employee hours, and other

pertinent topics. We have gleaned interesting data points from this survey. For example, 41.6 percent of small businesses estimate that more than 6 months will pass before their business returns to its normal level of operations – much longer than anticipated. We also saw that as heavy lockdowns were lifted, there was a surge in hiring, which then slowed and is now starting to rise again.

The Household Pulse Survey asks questions about changes in consumer habits, adherence to stay-at-home orders, intent to get a COVID-19 vaccination, whether masks are being worn, and pandemic impacts on mental health and school closures. This information has been exceptionally valuable for local officials and national-level decision makers responding to the pandemic. For example, data from the last two weeks in February show that one in three (34.6 percent) adults live in households that are having a somewhat or very difficult time paying their usual household expenses. And, among adults who have yet to receive a vaccine, more than half (52.5 percent) plan to definitely get one when available.

It's not enough to understand business trends or household impacts. We have to tie the data together in a way that helps us understand how communities as a whole are positioned to respond to extreme events, whether those be natural disasters like hurricanes and wildfires or pandemics. In response to the onset of COVID-19, we created a model that measured a community's ability to absorb and endure impacts from a disaster called the Community Resilience Estimates. We recognized instantly the value these estimates held for understanding communities' risk and elasticity in the face of the pandemic.

The idea behind the Community Resilience Estimates is that different communities will fare differently to extreme events depending on individual and household characteristics within the community. The estimates use 11 risk factors to model resilience, including: income to poverty ratio, number of adults in the household, number of people in a household, population density, education level, employment status, languages spoken, disability characteristics, health insurance status, age, and presence of high-risk health conditions. While other entities produce similar metrics, they are based on publicly available data. Our estimates are based on our full suite of information increasing the reliability of our numbers for areas not typically represented well in other data sets, for example rural areas.

Everyone responding to the COVID-19 pandemic is working in extraordinary conditions and circumstances. We want to provide as much information as we can, and to do so in a user-friendly way. To help everyone access all of these recent innovations as well as other relevant data sets, we created the COVID-19 Data Hub to centralize information pertinent to the pandemic. It is on our census.gov website and has a variety of interactive features that allow users to customize the data they need to inform a variety of COVID-19 response and recovery-related decisions.

Conclusion

We talk a lot about how the 2020 Census shapes the future for families, communities and our nation. But what we've seen over the past year is that we can produce information that tells us how to respond and actively shape our future in the moment. This is why we do our job. We are

mathematicians, statisticians, demographers, economists, geographers and computer specialists, and we love numbers. But more importantly, we appreciate what these numbers can do to help improve all of our lives. They matter to people every day, so we know that we must take the time to get it right so they are as effective as they can be. Getting it right for the decennial census meant delaying our original schedule so we could conduct it safely and taking the time needed to process the data so we can ensure we produce high quality statistics. Getting it right also means making sure that everyone feels safe responding to our censuses and surveys so that everyone in the country is reflected in the statistics we produce. People have to know that we will guard their privacy zealously if we want them to entrust us with their personal information.

Thank you allowing me the time to highlight the Census Bureau's impact on families, communities and the nation – today and as we shape our future together. Your support has been instrumental in conducting a decennial census in extraordinary circumstances. We are excited to engage more frequently with interested stakeholders around all of our censuses and surveys, but know there is particular interest in the 2020 Census results. I commit to you that we will redouble our efforts to explain our work so the public can be as confident as we are in the information we release.

I also want to reiterate that the Census Bureau only has one shot to count every person living in the country and we want it to be as complete and accurate as possible. Thank you and I look forward to taking any questions.